

On matrix polynomials with real roots

Leonid Gurvits*

Leiba Rodman†

Abstract

It is proved that the roots of combinations of matrix polynomials with real roots can be recast as eigenvalues of combinations of real symmetric matrices, under certain hypotheses. The proof is based on recent solution of the Lax conjecture. Several applications and corollaries, in particular concerning hyperbolic matrix polynomials, are presented.

Key Words. Hyperbolic polynomials, matrix polynomials.

2000 Mathematics Subject Classification. 47A56, 15A57.

1 Main result

A polynomial is called *hyperbolic* if all its roots are real. It is a classical well studied class of polynomials (see, e.g., [16]). There are at least two useful ways to extend this notion to polynomials with complex $n \times n$ matrix coefficients, in short, matrix polynomials. Thus, a monic (i.e., with leading coefficient I_n , the $n \times n$ identity matrix) matrix polynomial $L(z)$ of degree ℓ is said to be *hyperbolic*, if for every nonzero $x \in \mathbb{C}^n$, the n -dimensional vector space of columns with complex components, the polynomial equation

$$\langle L(z)x, x \rangle = 0 \tag{1.1}$$

has ℓ real roots (counted with multiplicities). We denote here by $\langle \cdot, \cdot \rangle$ the standard inner product in \mathbb{C}^n . An $n \times n$ monic matrix polynomial $L(z)$ of degree ℓ will be called *weakly hyperbolic* if $\det L(z) = 0$ has $n\ell$ real roots (multiplicities counted). Note that our terminology differs slightly from the terminology in some sources (for example, [14]). Clearly, every hyperbolic matrix polynomial is weakly hyperbolic, and the coefficients of every hyperbolic matrix polynomial are Hermitian matrices. See, e.g., [15, 1, 14, 10] for the theory and applications of hyperbolic matrix and operator polynomials.

*Los Alamos National Laboratory, Los Alamos, NM 87545, USA. E-mail: gurvits@lanl.gov.

†College of William and Mary, Department of Mathematics, P. O. Box 8795, Williamsburg, VA 23187-8795, USA. E-mail: lxrodm@math.wm.edu. The research leading to this article was done while the second author visited LANL, whose hospitality is gratefully acknowledged. The research of second author is partially supported by NSF Grant DMS-9988579.

In this note we prove the following theorem. It states that the roots of combinations of hyperbolic matrix polynomials can be recast as eigenvalues of combinations of real symmetric matrices, under certain hypotheses. We denote by \mathbb{R} the field of real numbers.

Theorem 1.1 *Let*

$$L(z) = \sum_{j=0}^{\ell} L_j z^j, \quad M(z) = \sum_{j=0}^{\ell} M_j z^j, \quad M_{\ell} = L_{\ell} = I,$$

be two monic $n \times n$ matrix polynomials such that

$$\alpha L(z) + (1 - \alpha)M(z) \quad \text{is weakly hyperbolic for every } \alpha \in \mathbb{R}. \quad (1.2)$$

Assume in addition that the $n \times n$ matrix $L_{\ell-1} - M_{\ell-1}$ has n real eigenvalues (counted with multiplicities). Then there exist $n\ell \times n\ell$ real symmetric matrices A and B such that for every $\alpha \in \mathbb{R}$, the roots of $\det(\alpha L(z) + (1 - \alpha)M(z))$, counted according to their multiplicities, coincide with the eigenvalues of $\alpha A + (1 - \alpha)B$, also counted according to their multiplicities.

Conversely, if the roots of $\det(\alpha L(z) + (1 - \alpha)M(z))$, coincide with the eigenvalues of $\alpha A + (1 - \alpha)B$ (counted with multiplicities) for every $\alpha \in \mathbb{R}$, where A and B are fixed real symmetric $n\ell \times n\ell$ matrices, then (1.2) holds and all eigenvalues of $L_{\ell-1} - M_{\ell-1}$ are real.

Proof. Let

$$C_L = \begin{pmatrix} 0 & I_n & 0 & \dots & 0 \\ 0 & 0 & I_n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_n \\ -L_0 & -L_1 & -L_2 & \dots & -L_{\ell-1} \end{pmatrix}$$

and

$$C_M = \begin{pmatrix} 0 & I_n & 0 & \dots & 0 \\ 0 & 0 & I_n & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I_n \\ -M_0 & -M_1 & -M_2 & \dots & -M_{\ell-1} \end{pmatrix}$$

be the companion matrices of $L(z)$ and of $M(z)$, respectively. Then $\alpha C_L + (1 - \alpha)C_M$ is the companion matrix of $\alpha L(z) + (1 - \alpha)M(z)$, and therefore the roots of $\det(\alpha L(z) + (1 - \alpha)M(z))$ (counted with multiplicities) coincide with the eigenvalues of $\alpha C_L + (1 - \alpha)C_M$ (also counted with multiplicities), for every $\alpha \in \mathbb{R}$. (This is a standard fact in the theory of matrix polynomials, see, for example, [5].) Thus, $\alpha C_L + (1 - \alpha)C_M$ has

$n\ell$ real eigenvalues for every $\alpha \in \mathbb{R}$. Consider the homogeneous polynomial of three real variables α, β, γ :

$$P(\alpha, \beta, \gamma) := \det(\alpha C_L + \beta C_M - \gamma I_{n\ell}).$$

If $\alpha + \beta \neq 0$, the polynomial $P(\alpha, \beta, \gamma)$ (as a polynomial of γ) has $n\ell$ real roots (counted with multiplicities). If $\alpha + \beta = 0$, then

$$P(\alpha, \beta, \gamma) = \pm \gamma^{n(\ell-1)} \cdot \det(\alpha L_{\ell-1} + \beta M_{\ell-1} - \gamma I_n)$$

also has $n\ell$ real roots, by hypothesis of the theorem. Thus, $P(\alpha, \beta, \gamma)$ is hyperbolic in the direction of $(0, 0, 1)$, in the sense of the Lax conjecture, see [11, 13]. By the main result of [13] (the proof in [13] is based on [7, 17]), we have

$$P(\alpha, \beta, \gamma) = \det(\alpha A + \beta B - \gamma I)$$

for some real symmetric matrices A and B . The direct statement of the theorem follows.

To prove the converse statement, simply reverse the argument, taking into account that real symmetric matrices have all eigenvalues real. \square

For further development of the theory of hyperbolic polynomials of several variables and many applications, in particular, mixed determinants, see [6].

2 Corollaries and applications

We start by recalling Obreschkoff's theorem (see [16, 3]), which will be needed in the proof of the next corollary: Two real scalar polynomials $f(z)$ and $g(z)$ of degrees ℓ and $\ell - 1$, respectively, have the property that $f(z) + tg(z)$ has ℓ real roots (counted with multiplicities) for every real t if and only if $f(z)$ and $g(z)$ have ℓ and $\ell - 1$ real roots, respectively, and the roots of $f(z)$ and of $g(z)$ interlace (the cases when $f(z)$ or $g(z)$ have multiple roots and/or when $f(z)$ and $g(z)$ have common roots are not excluded here).

A proof of Obreschkoff's theorem can be given using the approach of Theorem 1.1, as follows (below, we formulate Obreschkoff's theorem in a slightly different but equivalent form):

Proposition 2.1 *The following statements are equivalent for scalar distinct monic relatively prime polynomials $f(z)$ and $h(z)$ of degree ℓ :*

- (1) *The polynomials $\alpha f(z) + \beta h(z)$, $\alpha, \beta \in \mathbb{R}$, $\alpha^2 + \beta^2 \neq 0$, have all their roots real;*
- (2) *The polynomials $\alpha f(z) + (1 - \alpha)h(z)$, $\alpha \in \mathbb{R}$, have all their roots real;*

(3) $f(z)$ has all its roots real, and the quotient $h(z)/f(z)$ has the form

$$\frac{h(z)}{f(z)} = 1 + \sum_{j=1}^p \frac{c_j}{z - \lambda_j}, \quad (2.1)$$

where $\lambda_j \in \mathbb{R}$, and the real numbers c_j are all of the same sign;

(4) Both $f(z)$ and $h(z)$ have ℓ distinct real roots and the roots of $f(z)$ and of $h(z)$ interlace.

Proof. (1) clearly implies (2) . (3) and (4) are equivalent : Indeed, if (3) or (4) holds true, then $f(z)$ has necessarily simple roots, and denoting the roots of $f(z)$ by $\lambda_1 < \dots < \lambda_p$, we see that in the representation (2.1),

$$\text{sign}(c_j) = \text{sign} \frac{h(z)}{f(z)} \quad \text{as } z \rightarrow \lambda_j, z > \lambda_j,$$

whereas

$$-\text{sign}(c_{j+1}) = \text{sign} \frac{h(z)}{f(z)} \quad \text{as } z \rightarrow \lambda_{j+1}, z < \lambda_{j+1}.$$

These equalities imply that the roots of $f(z)$ and $h(z)$ interlace if and only if all the c_j 's are of the same sign.

We next prove that (2) implies (3). Arguing as in the proof of Theorem 1.1, we obtain that the characteristic polynomial of $\alpha C_f + \beta C_h$ coincides with the characteristic polynomial of $\alpha A + \beta B$, for all $\alpha, \beta \in \mathbb{R}$, where A and B are fixed (independent of α and β) distinct real symmetric matrices. Taking $\alpha - \beta = 0$ we see that $\text{rank}(A - B) \leq 1$. Since polynomials f and h are distinct it follow that $A = B \pm xx^T$ for some nonzero vector x . Now

$$\begin{aligned} f(z) &= \det(zI - C_f) = \det(zI - A), \\ h(z) &= \det(zI - C_h) = \det(zI - B), \end{aligned}$$

and

$$\frac{h(z)}{f(z)} = \det((zI - B)(zI - A)^{-1}) = \det(I \pm xx^T(zI - A)^{-1}) = 1 \pm x^T(zI - A)^{-1}x.$$

This reduces, upon applying a diagonalizing real orthogonal transformation $A \mapsto U^T A U$, and replacing x with $U^T x$, to (2.1) with the real numbers c_j of the same sign, as required.

Finally , let us prove the implication (3) \implies (1) . This means to prove that if (3) holds then for any real γ the equation $\frac{h(z)}{f(z)} + \gamma = 0$ does not have roots with nonzero imaginary part .

Consider a complex number $z = a + bi$, its real part $Re(z) = a$, its imaginary part $Im(z) = b$. If (3) holds then

$$Im\left(\frac{h(z)}{f(z)}\right) = Im\left(\sum_{j=1}^p \frac{c_j}{z - \lambda_j}\right),$$

where λ_j are real and the real numbers c_j are all of the same sign. Assume wlog that all c_j are positive. As

$$Im((z - \lambda_j)^{-1}) = \frac{-b}{(a - \lambda_j)^2 + b^2},$$

thus we get that $Im\left(\frac{h(z)}{f(z)}\right) = -b \sum_{j=1}^p \frac{c_j}{(a - \lambda_j)^2 + b^2}$.

Therefore $Im\left(\frac{h(z)}{f(z)}\right) \neq 0$ if $Im(z) \neq 0$. This means that the equation $\frac{h(z)}{f(z)} + \gamma = 0$ does not have roots with nonzero imaginary part for all real γ .

□

We observe that checking condition (3) can be conveniently done using semidefinite programming. Indeed, let $h(z)$ and $f(z)$ be monic scalar polynomials with $f(z)$ having all roots real, and consider a minimal realization

$$\frac{h(z)}{f(z)} = 1 + \tilde{C}(zI - \tilde{A})^{-1}\tilde{B},$$

where \tilde{C} , \tilde{A} , and \tilde{B} are real matrices. It is easy to see, using the uniqueness of a minimal realization up to a state isomorphism (similarity), that (3) holds, with the c_j 's positive, if and only if there exists a positive definite matrix P such that

$$\tilde{A}P = P\tilde{A}^T, \quad P\tilde{C}^T = \tilde{B}.$$

The latter problem is a semidefinite programming problem.

Another equivalent semidefinite programming problem is based on the following nice reformulation of Proposition 2.1 :

The conditions of Proposition 2.1 are equivalent to the existence of nonsingular real matrix D such that both $DC_f D^{-1}$ and $DC_h D^{-1}$ are real symmetric.

(Our proof of this statement is essentially the same rank one perturbation argument as in the proof of Proposition 2.1).

This gives the following semidefinite programming problem :
is there exists a real positive definite $P \succ 0$ such that

$$PC_f = C_f^T P \text{ and } PC_h = C_h^T P$$

Our next corollary involves hyperbolic matrix polynomials.

Corollary 2.2 *Let $L(z)$ be a hyperbolic matrix polynomial. Then there exist $n\ell \times n\ell$ real symmetric matrices A and B such that the roots of $\det(L(z) + tL'(z))$ coincide with the eigenvalues of $A + tB$ (multiplicities counted), for every real number t . Here, $L'(z)$ is the derivative of $L(z)$ with respect to z .*

Proof. . By Obreschkoff's theorem, the matrix polynomial $L(z) + tL'(z)$ is hyperbolic for every real t . Now apply Theorem 1.1 with $M(z) = L(z) + L'(z)$. \square

Note that the condition (1.2) implies (but is not equivalent to) the condition that every convex combination of $L(z)$ and $M(z)$ is weakly hyperbolic. It turns out that the latter condition can be conveniently expressed for hyperbolic matrix polynomials, which we will do next.

Let $L(\lambda)$ be a hyperbolic $n \times n$ matrix polynomial. For every $x \in \mathbb{C}^n$, $\|x\| = 1$, let

$$\lambda_1(x) \leq \lambda_2(x) \leq \cdots \leq \lambda_\ell(x)$$

be the roots of equation (1.1) arranged in the nondecreasing order. The sets

$$\Delta_j(L) := \{ \lambda_j(x) \mid x \in \mathbb{C}^n, \|x\| = 1 \},$$

called the *spectral zones* of $L(\lambda)$, are obviously closed intervals on the real line:

$$\Delta_j(L) = [\delta_j^-(L), \delta_j^+(L)], \quad j = 1, 2, \dots, \ell.$$

A basic result in the theory of hyperbolic matrix and operator polynomials ([14, Theorem 31.5], for example), states that two spectral zones either are disjoint, or have only one point in common.

Proposition 2.3 *Let $L(\lambda)$ and $M(\lambda)$ be two hyperbolic matrix polynomials of degree ℓ . Then every convex combination $\alpha L(z) + (1 - \alpha)M(z)$, $0 \leq \alpha \leq 1$, is hyperbolic if and only if their spectral zones satisfy the inequalities*

$$\max\{\delta_j^+(L), \delta_j^+(M)\} \leq \min\{\delta_{j+1}^-(L), \delta_{j+1}^-(M)\}, \quad j = 1, \dots, \ell - 1.$$

For the proof apply [3, Theorem 2.1]; this theorem gives necessary and sufficient conditions for all linear combinations of two given scalar polynomials to be hyperbolic.

Using Theorem 1.1 and inequalities for eigenvalues of real symmetric matrices (see for example [12]), one can derive inequalities for eigenvalues of weakly hyperbolic matrix polynomials. We illustrate this for the case of the Horn inequalities. For a Hermitian $m \times m$ matrix X , we write its eigenvalues (repeated according to their multiplicities) in a non-decreasing order:

$$\lambda_1(X) \leq \lambda_2(X) \leq \cdots \leq \lambda_m(X).$$

An ordered triple (U, S, T) of nonempty subsets of $\{1, 2, \dots, m\}$ is said to be a *Horn triple* (with respect to m) if the cardinalities of U , S , and T are the same, and the *Horn inequalities*

$$\sum_{i \in U} \lambda_i(X + Y) \leq \sum_{j \in S} \lambda_j(X) + \sum_{k \in T} \lambda_k(Y)$$

hold true for every pair of Hermitian $m \times m$ matrices X and Y . A description of all Horn triples is known [8, 9]; see also the surveys [4, 2]. For a weakly hyperbolic $n \times n$ matrix polynomial $L(z)$ of degree ℓ , we arrange the roots of $\det(L(z))$ in the non-decreasing order:

$$d_1(L) \leq d_2(L) \leq \dots \leq d_{n\ell}(L).$$

Let $T = \{1 \leq i_1 < i_2 < \dots < i_m \leq n\ell\} \subset \{1, 2, \dots, n\ell\}$, define $\bar{T} = \{n\ell - i_m < n\ell - i_{m-1} < \dots < n\ell - i_1\}$.

Theorem 2.4 *Let $L(z)$ and $M(z)$ be monic $n \times n$ matrix polynomials satisfying the hypotheses of Theorem 1.1. Then for every Horn triple (U, S, T) with respect to $n\ell$, and for every $\alpha \in \mathbb{R}$, the inequality*

$$\sum_{i \in U} d_i(\alpha L + (1 - \alpha)M) \leq \alpha \left(\sum_{j \in S_\alpha} d_j(L) \right) + (1 - \alpha) \left(\sum_{k \in T_{1-\alpha}} d_k(M) \right)$$

holds true. Here $S_\alpha = S$ if $\alpha \geq 0$ and $S_\alpha = \bar{S}$ if $\alpha < 0$.

Proof. Let A and B be as in Theorem 1.1. Then we have, using Theorem 1.1 and the Horn inequalities:

$$\begin{aligned} \sum_{i \in U} d_i(\alpha L + (1 - \alpha)M) &= \sum_{i \in U} \lambda_i(\alpha A + (1 - \alpha)B) \\ &\leq \sum_{j \in S} \lambda_j(\alpha A) + \sum_{k \in T} \lambda_k((1 - \alpha)B) \\ &= \alpha \left(\sum_{j \in S_\alpha} \lambda_j(A) \right) + (1 - \alpha) \left(\sum_{k \in T_{1-\alpha}} \lambda_k(B) \right) \\ &= \alpha \left(\sum_{j \in S_\alpha} d_j(L) \right) + (1 - \alpha) \left(\sum_{k \in T_{1-\alpha}} d_k(M) \right), \end{aligned}$$

and the proof is complete.

References

- [1] L. Barkwell, P. Lancaster, and A. S. Markus. Gyroscopically stabilized systems: a class of quadratic eigenvalue problems with real spectrum. *Canad. J. Math.* 44 (1992), 42–53.

- [2] R. Bhatia. Linear algebra to quantum cohomology: the story of Alfred Horn's inequalities. *Amer. Math. Monthly* 108 (2001), 289–318.
- [3] J. P. Dedieu. Obreschkoff's theorem revisited: What convex sets are contained in the set of hyperbolic polynomials? *Journal of Pure and Applied Algebra* 81 (1992), 269–278.
- [4] W. Fulton. Eigenvalues, invariant factors, highest weights, and Schubert calculus. *Bull. Amer. Math. Soc. (N.S.)* 37 (2000), 209–249.
- [5] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials*. Academic Press, 1982.
- [6] L. Gurvits. Combinatorics hidden in hyperbolic polynomials and related topics, preprint (2004), available from xxx.lanl.gov.
- [7] J. W. Helton and V. Vinnikov. Linear matrix inequality representation of sets, preprint (2003).
- [8] A. A. Klyachko. Stable bundles, representation theory, and Hermitian operators. *Selecta Math. (N. S.)* 3 (1998), 419–445.
- [9] A. A. Klaychko. Vector bundles, linear representations, and spectral problems. *Proceedings of the International Congress of Mathematicians*. Vol. II 599–613, Higher Ed. Press, Beijing, 2002.
- [10] P. Lancaster, A. S. Markus, and V. I. Matsaev. Definitizable operators and quasi-hyperbolic operator polynomials. *J. Funct. Anal.* 131 (1995), 1–28.
- [11] P. D. Lax, Differential equations, difference equations and matrix theory. *Comm. Pure Appl. Math.* 11 (1958), 175–194.
- [12] A. S. Lewis, and M. L. Overton. Eigenvalue optimization. *Acta Numerica* (1996), 149 - 190.
- [13] A. S. Lewis, P. A. Parrilo, and M. V. Ramana. The Lax conjecture is true, preprint (2003).
- [14] A. S. Markus. *Introduction to the spectral theory of polynomial operator pencils*. Translations of Mathematical Monographs, 71. Amer. Math. Soc., Providence, RI, 1988.
- [15] A. S. Markus, V. I. Macaev, and G. I. Russu. Certain generalizations of the theory of strongly damped pencils to the case of pencils of arbitrary order. *Acta Sci. Math. (Szeged)* 34 (1973), 245–271.

- [16] N. Obreschkoff. *Verteilung und Berechnung der Nullstellen reellen Polynome*. VEB Deutscher Verlag der Wissenschaften, 1963.
- [17] V. Vinnikov, Selfadjoint determinantal representations of real plane curves. *Math. Ann.* 296 (1993), 453–479.